

Reproductive compensation and human genetic disease

IAN M. HASTINGS*

Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK

(Received 5 October 2000 and in revised form 5 December 2000)

Summary

The effects of reproductive compensation on the population genetics of sex-linked recessive lethal mutations are investigated. Simple equations are presented which describe these effects, and so complement existing population genetic theory. More importantly, this type of mutation is responsible for several severe human genetic diseases such as Duchenne muscular dystrophy. It is argued that the applications of three modern reproductive technologies – effective family planning, *in utero* diagnosis with termination, and embryo sexing – will lead to reproductive compensation. The adoption of any of these technologies may rapidly elevate the frequencies of those mutations which are lethal in childhood. This increase is large, in the order of 33% upwards, and occurs rapidly over two to five generations. It also depends on the source of mutations, the effect being larger if most mutations are paternal. *In utero* diagnosis and/or embryo sexing increase the frequency of the mutation, but simultaneously decrease disease incidence by preventing the birth of affected offspring. In contrast, effective family planning may rapidly increase both mutation frequency and disease incidence.

1. Introduction

It is widely assumed that genetic diseases persist in human populations as a balance between mutation (which creates the mutant genes) and natural selection (which eliminates them through the death of affected individuals): the ‘mutation/selection balance’. Standard population genetic analysis of mutation/selection balance implicitly assumes that once fertilization of gametes has occurred, the subsequent fates of the parents and offspring are independent. This has the advantage of generality across species but can lead to errors, particularly when applied to mammals. Taking humans as an example: a human mother enters a period of amenorrhoea following conception of a child and cannot reconceive during this period. This period may last from 9 months (i.e. until immediately after birth) in modern societies to 4 years in hunter/gatherer societies (Short, 1994). If a human embryo or child dies early in development, the mother will immediately re-enter oestrus and can conceive another child. In effect, the dead child has been partially replaced, an effect known as ‘reproductive

compensation’. The effect is most noticeable in humans because we generally conceive a single embryo and have a long gestation period. However, it also occurs in many other situations (Hastings, 2000): for example, many bird species produce an ‘extra’ hatchling which can only survive if one of its older siblings dies. The consequences of reproductive compensation are sporadically considered in the literature (a few representative examples being Race (1944), Hamilton (1966) and Charlesworth (1994)) but in general are not considered in any systematic way in the mainstream literature (it is, for example, absent from the textbooks by Falconer & Mackay (1996), Hartl & Clark (1997) and Maynard Smith (1998)), despite having potentially large effects on equilibrium gene frequencies (Hastings, 2000).

Technological changes in human reproduction can result in reproductive compensation and alter the dynamics of genetic diseases away from that described by standard population genetic theory. Consider the following example: a family planned to have three children but one subsequently died from a genetic disease. If a further child was produced to bring family size back up to three, then full reproductive compensation occurred. This is, of course, a very

* Tel. +44 (0) 151 708 9393. Fax: +44 (0) 151 708 8733.
e-mail: hastings@liverpool.ac.uk.

inflexible model of human behaviour (discussed later) but clearly shows that effective family planning can alter the population genetics of the disease. A second technology resulting in reproductive compensation is the advent of techniques to identify and terminate affected individuals *in utero* (Motulsky, 1997; Milunsky, 1998). Amniocentesis (the extraction of fetal cells from the amniotic sac) and choriovillous sampling (the direct sampling of fetal tissue) are routinely offered to pregnant women in the United Kingdom and allow genetic screening of the fetus as early as 18 and 12 weeks, respectively, after conception. It seems likely that a mother choosing to undergo these tests would choose to terminate a fetus with a major genetic disease, and continue to reconceive until an unaffected fetus is produced. This constitutes reproductive compensation, again affecting the dynamics of genetic diseases that kill during childhood. A third technology is being developed to reduce the incidence of sex-linked recessive lethal disease. These affect only boys so the intention is to sex the embryos of mothers known to carry the disease and remove male fetuses to ensure that only daughters are produced. This has ethical problems because, in the absence of a reliable genetic marker, 50% of discarded male fetuses will be healthy. A alternative currently under development is to eliminate Y-bearing sperm *in vitro* prior to insemination. This is not reliable at present but much work is being undertaken on animal models and we can anticipate the development of an easy and effective technology to achieve this and examine its consequences. As noted by Emery (1993), this technology effectively replaces sons with carrier daughters and the frequency of the disease gene will increase, the magnitude and speed of the process being unknown at present.

Hastings (2000) briefly noted that human reproductive technology could result in reproductive compensation, and examined its effects on autosomal genetic disease. The present study extends this work to explicitly examine the effects of differing reproductive technologies, and to investigate the consequences for sex-linked, rather than autosomal, diseases. Many of the more common and debilitating human genetic diseases are caused by sex-linked recessive lethal mutations, familiar examples being Duchenne muscular dystrophy which affects 1 in 3500 live male births, and Alport syndrome and haemophilia A, each of which affect 1 in 5000 male births (Conner & Ferguson-Smith, 1997). A secondary objective is that of theoretical completeness. Mutations which kill very early in development will follow dynamics determined by reproductive compensation, irrespective of maternal reproductive behaviour (Hastings, 2000). This includes all genes which prevent early division of the zygote, which presumably includes all cell cycle, DNA/RNA synthesizing, and possibly many other

'housekeeping', genes. Differing patterns of molecular variation at different loci can be used to infer differing selective processes (such as selective sweeps or balancing selection), but to my knowledge, the effects of differential selection pressures attributable to the presence or absence of reproductive compensation have not been explicitly considered. It will be shown that a similar pattern emerges as for autosomal genes: that equilibrium mutation frequencies at sex-linked loci subject to reproductive compensation may be much higher than at those loci not subject to reproductive compensation. Furthermore, the extent of this difference is determined by the relative mutation rates in the two sexes.

2. Methods

We follow the method and notation of Crow (1986) to track the frequency of the heterozygote (carrier) females, H , and the frequency of the mutant gene in the population, q . Since sons inherit only a single copy of the X chromosome, q is also the frequency of affected male genotypes at conception. The normal wild-type gene is designated $+$, and the mutant gene as a . We also assume that the mutation is fully recessive so there is no natural selection against female heterozygotes; this assumption will be relaxed later. The standard theory on sex-linked recessive lethals is straightforward. The recurrence equation for H is:

$$H' = \frac{H}{2} + \frac{H\mu_f}{2} + (1-H)\mu_f + \mu_m,$$

where μ_f and μ_m are mutation rates in females and males respectively. The products $H\mu_f$ and $H\mu_m$ can be regarded as negligible so we obtain:

$$H' \approx \frac{H}{2} + \mu_f + \mu_m. \quad (1)$$

Since the male phenotype is lethal, the allele can only be passed on through the female line, so

$$\begin{aligned} q' &= \frac{H}{2} + \frac{H\mu_f}{2} + (1-H)\mu_f \\ &\approx \frac{H}{2} + \mu_f. \end{aligned} \quad (2)$$

At equilibrium $H' = H$ so the equilibrium frequency of carriers \hat{H} is $2\mu_f + 2\mu_m$ (from equation 1), giving the standard equilibrium frequency of q (from equation 2) as

$$\hat{q} = 2\mu_f + \mu_m. \quad (3)$$

(i) Full reproductive compensation

The reproductive output from carrier $+a$ females in the standard model is 1/4 each of the following four

genotypes: a males, $+$ males, $++$ females and $+$ a females. With full reproductive compensation, the reproductive output from $+$ a females is $1/3$ each of $+$ males, $++$ females and $+$ a females. One method of investigating reproductive compensation is to incorporate these changes in offspring frequency (Hastings, 2000). A more elegant means of achieving the same result for sex-linked lethals is simply to note that the relative fecundity of carriers, in terms of female offspring, is increased from 1 to $4/3$ (because the proportion of female offspring produced by carriers increases from $1/2$ to $2/3$, i.e. a ratio of $1:4/3$) Thus under a model of full reproductive compensation, equation (1) is inaccurate and should be replaced by:

$$H' \approx \frac{H(4/3)}{2} + \mu_f + \mu_m,$$

so

$$\hat{H} = 3\mu_f + 3\mu_m$$

and application of equation (2) gives

$$\hat{q} = 2.5\mu_f + 1.5\mu_m. \quad (4)$$

(H is not scaled by $4/3$ in equation (2) so equation (4) predicts gene frequency prior to selection and reproductive compensation.)

(ii) Embryo sexing

Assume a proportion z of carrier females can be identified prior to reproduction and subsequently manipulated to produce only daughters. The contribution of carrier females to the daughters of their generation is therefore increased to $(1+z)$, provided H is small enough that (i) the normalizing factor $H(1+z) + (1-H) \approx 1$ and (ii) the overall surplus of females in the population is small. Equation (1) then becomes:

$$H' \approx \frac{H}{2} (1+z) + \mu_f + \mu_m,$$

giving

$$\hat{H} = \frac{2\mu_f + 2\mu_m}{1-z}$$

so that

$$\hat{q} = \frac{2\mu_f}{1-z} + \mu_m. \quad (5)$$

The recurrence equation is initially correct but the equilibrium equations break down as z nears unity because there is no direct selection against the mutation, only indirect selection on the sex ratio, so H becomes large and assumptions (i) and (ii) detailed

above, and the assumption that $H\mu$ can be regarded as negligible, are violated.

(iii) Effects of reproductive restraint or natural selection against the female carrier

Both the standard model (equation 3) and one assuming full reproductive compensation (equation 4) rest on a somewhat inflexible view of human nature. The standard model assumes females continue to reproduce as fast as possible with their reproductive output completely unaffected by the birth of an affected son. Similarly, the model of full reproductive compensation rests on the assumption that if a mother has decided to have, for example, three offspring, then affected offspring are discounted and reproduction continues until three healthy offspring have been born. In reality, the time and emotional investment required to care for an affected offspring will have dissuaded many mothers from having subsequent children. Similarly, the risk (1 in 4, assuming it was not a *de novo* mutation) of producing another affected child will have been a disincentive. Both factors are likely to reduce family size. It is possible to incorporate the effects of this reproductive restraint, r , where r is the reduction in family size of carriers. In the standard model, equation (1) becomes:

$$H' \approx \frac{H}{2} (1-r) + \mu_f + \mu_m,$$

giving

$$\hat{H} = \frac{2\mu_f + 2\mu_m}{1+r},$$

so that

$$\hat{q} = \frac{\mu_f + \mu_m}{1+r} + \mu_r. \quad (6)$$

Similarly, in the model of reproductive compensation, equation (1) becomes:

$$H' \approx \frac{H(4/3)(1-r)}{2} + \mu_f + \mu_m,$$

giving

$$\hat{H} = \frac{3\mu_f + 3\mu_m}{1+2r},$$

so that

$$\hat{q} = \frac{1.5\mu_f + 1.5\mu_m}{1+2r} + \mu_r. \quad (7)$$

The effects of natural selection against the female heterozygote are mathematically identical to those of reproductive restraint. In the former their reproductive

output is reduced by a factor $(1-s)$ and in the latter by $(1-r)$, where s is the strength of natural selection. Substituting s for r in equations (6) and (7) therefore allows the investigation of mutations where natural selection does appear to act against heterozygous females, such as ornithine transcarboxylase deficiency (which is lethal in male neonates and has variable penetrance in females depending on their pattern of X chromosome inactivation in the liver).

3. Results

Equations (3) and (4) show that the frequency of sex-linked lethals will rise as a consequence of reproductive compensation. Interestingly, the magnitude of this increase depends on the source of the mutations. If mutation rates are the same in both sexes, the equilibrium frequency rises from 3μ to 4μ , i.e. an increase of 33%. If mutation occurs predominantly in males ($\mu_f \approx 0$) then equilibrium frequencies rise from μ_m to $1.5\mu_m$, an increase of 50%. Similarly, equations (3) and (5) show that embryo sexing can greatly increase the frequency of the mutation in the population. For example, if a fifth of carriers are identified prior to reproduction ($z = 0.2$) there will be an increase in frequency of 16% (if mutation rate is the same in both sexes) or up to 25% (as mutation becomes predominantly male-driven). If 50% of carriers can be identified ($z = 0.5$) then the corresponding figures are 66% and 100%. Thus either reproductive compensation or embryo sexing may dramatically increase the frequency of the mutation in the population. Furthermore, this increase will be

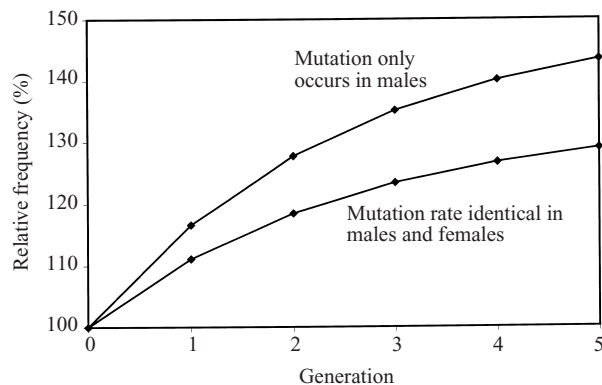


Fig. 1. The predicted rise in frequency of a sex-linked recessive lethal mutation following the advent of reproductive compensation. Mutation is assumed to occur either only in males, or in both males and females at an equal rate. This figure was produced by taking the equilibrium values for mutation and carrier frequency from the standard formula (equation 3) as the starting point, then using the recurrence formulae used to derive equation (4) to follow the effects of the introduction of reproductive compensation; the curves asymptotically approach the equilibrium given in equation (4).

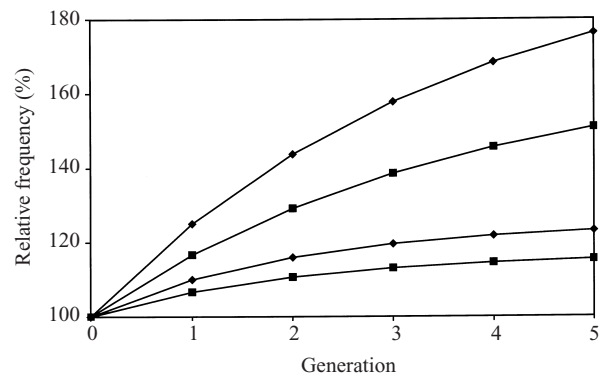


Fig. 2. The predicted rise in frequency of a sex-linked recessive lethal mutation as a consequence of embryo sexing. The proportion of carriers identified (z in equation 5) was assumed to be either 0.2 (the lower pair of lines) or 0.5 (the upper pair of lines). Within these pairs, the upper lines represents the boundary when all mutations occur in males, while the lower lines represent the case of equal mutation rates in both sexes. This figure was produced in an analogous manner to that used in Fig. 1: equilibrium values for mutation and carrier frequency from the standard formula (equation 3) provided the starting point and recurrence formulae from the derivation of equation (5) were used to follow the consequences of introducing embryo sexing; the curves asymptotically approach the equilibrium given in equation (5).

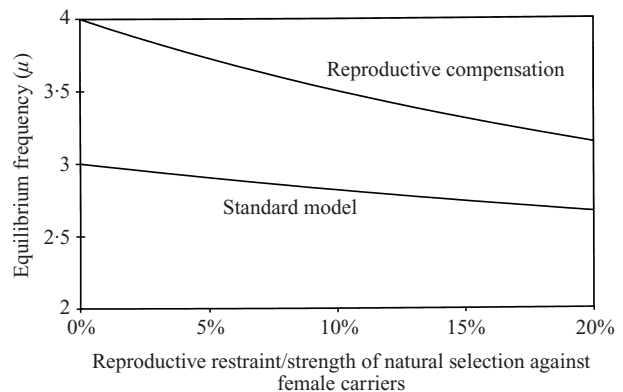


Fig. 3. The effects of reproductive restraint (or, alternatively, the strength of natural selection against the female carrier) on the equilibrium frequencies of a sex-linked recessive lethal mutation. The equilibrium frequencies based on models of full reproductive compensation (equation 7) or the standard model (equation 6) are shown. This figure was based on the assumption of equal mutation rates in the two sexes, but qualitatively similar results are obtained if mutation is restricted to males (data not shown).

extremely rapid: with full reproductive compensation and equal mutation rates in the two sexes, there is an 11% increase in the first generation and a further 7% in the second, the frequency rising asymptotically to 133%. If mutation occurs overwhelmingly in males there is a 17% increase in the first generation and a

10% increase in the second, again rising asymptotically to 150% (Fig. 1). Similar results are obtained as a consequence of embryo sexing (Fig. 2): the magnitude of increase is large and the same very rapid initial increase occurs.

The effects of reproductive restraint are plotted in Fig. 3 assuming equal mutation rates in the sexes (the same qualitative pattern is noted if mutation occurs only in males). As the magnitude of reproductive restraint increases, the equilibrium frequency of the syndrome decreases and the difference between the predictions of the standard and reproductive compensation models is reduced.

4. Discussion

Reproductive compensation can result in much higher equilibrium frequencies of sex-linked recessive lethals than predicted by standard theory (equation 4 vs 3). It therefore seems inevitable that sex-linked loci whose products affect early development (which presumably include the housekeeping genes responsible for basic metabolism and cell division) will exhibit higher equilibrium frequency of lethals than those loci whose lethality acts later in life. Given the genetic technology now available in mouse and humans, and the discrete nature of the X chromosome, it seems plausible to test this hypothesis and obtain an estimate of the importance of reproductive compensation in mammalian population genetics. Standard genetic theory is based on the assumption that either reproductive compensation is absent (as may well be the case in many 'lower' organisms with external fertilization, no parental care and no sib competition), or that genetic lethality always occurs after the opportunity for compensation has elapsed. Neither assumption seems reasonable in mammals and it appears from these and earlier results (Hastings, 2000) that this may be a significant omission in standard population genetic theory.

Of more immediate social interest is the fact that modern reproductive technologies will result in reproductive compensation and have an impact on the frequencies of sex-linked lethal diseases. There are three features that characterize the change in frequencies following the adoption of reproductive technologies: it is large, it is rapid, and it is affected by source (maternal and/or paternal) of mutations. There are simple intuitive explanations for each of these three features. First, the increase occurs because in the standard model an affected male dies before reproducing and his death removes one mutated X chromosome from the population. With reproductive compensation the affected male will be replaced by an unaffected offspring, i.e. with an unaffected brother, a ++ sister or a +a sister, each with frequency 1/3. Thus there is a 1/3 chance he will be replaced by a

sister carrying the mutation so his death effectively only removes 2/3 of a gene from the population. Reproductive compensation effectively reduces selection against the mutation and its frequency rises in the population. Second, the frequency of the mutation changes rapidly because a mutated gene is eliminated as soon as it is passed onto a son. Two-thirds of the X chromosomes are in females and one-third in males; this means that the average 'lifespan' of the mutation is about three generations and the increase occurs over this timescale of a few generation (Figs. 1, 2). The speed of increase is the chief difference between diseases caused by sex-linked and autosomal recessive lethals. The latter will increase by around 45% as a consequence of the modern reproductive technology (Hastings, 2000) but their rate of increase will be much slower, initially in the region of 0.6% per generation; their slow rate of increase occurs because they are only rarely inherited with another defective copy of the gene and hence rarely lethal, their 'lifespan' in the population is much longer and the increase in frequency occurs on a timescale of hundreds of generations. Third, the effect is larger when most mutations occur in males because if an affected son arises through a spontaneous *de novo* mutation in his mother, then his death will not be replaced by a carrier sister. Conversely, if the son arises from a carrier female then, as noted above, there is a 1/3 chance he will be replaced by a carrier sister. Thus the greater the proportion of affected males produced from carrier females, the greater the effect of reproductive compensation. Equation (2) shows that as the spontaneous female mutation rate falls, then a greater proportion of affected sons are derived from carrier females compared with *de novo* maternal mutations. Reproductive compensation therefore has a more marked effect as female mutation rate declines.

The difference between the equilibrium frequencies predicted by the standard and reproductive-compensation models becomes smaller as the magnitude of reproductive restraint increases. The increase shown in Fig. 1 needs to be interpreted with this in mind. If reproductive restraint is present (and it seems intuitively likely) the increase predicted as a consequence of reproductive compensation may be an overestimate (although it will still be rapid). Alternatively, it may be that the present magnitude of restraint is much less than in the past as the application of genetic technology removes both the necessity of caring for affected children and the risk of producing an affected child. Thus the increase shown in Fig. 1 may actually be an underestimate as the starting frequency in real human populations may well have been significantly below that taken as the baseline (equation 3, rather than equation 6). It is impossible to draw any firm conclusions as the value of r will be specific to individual diseases and societies, depending

on the severity of the symptoms, age of onset, length of chronic symptoms, degree of familial support, and so on. It is, however, important to note that the frequency predicted by the standard model (equation 3) is an overestimate as it ignores reproductive restraint. The results therefore depend on female behaviour (reproductive restraint) but are independent of social structure: fathers always carry a single wild-type allele (since the mutant is male-lethal) so it is immaterial whether females conceive within a monogamous or a promiscuous society.

Adherence to a predetermined family size is the factor that leads to reproductive compensation and ultimately to an increase in disease incidence. However, strict adherence to a planned number of children was historically problematic for both practical reasons (in the absence of effective methods of contraception) and social reasons. Family planning in Western societies became much more effective following the social changes and introduction of the oral contraceptive pill in the 1960s. Forty years is a reasonable period given the timescale of the process (Fig. 1) so it is reasonable to ask whether an increase in sex-linked genetic diseases has occurred recently. This is not as straightforward as might be supposed as their prevalence must be ascertained by actively seeking out affected individuals (the syndromes are not obvious at birth so cannot be recorded at this point, and death certificates invariably record the proximal cause of death, such as heart failure). A recent comprehensive study in Northern Ireland by Hughes *et al.* (1996) reported a frequency of Duchenne muscular dystrophy (DMD) at the high end of the range of past reported incidences, but this may have been due to highly effective case detection rather than a real increase over time. Another confounding factor is that syndromes with similar symptoms may previously have been indistinguishable, even though their genetic basis is different. Current molecular technology allows these syndromes to be distinguished. Thus Emery (1991) in a large survey of DMD incidence noted that any increase over time may have been 'counterbalanced ... by improved ascertainment of true cases' and gone unnoticed. However, both the mathematics and logic suggest that such an increase is inevitable.

The consequences of genetic technology have been discussed in terms of the frequency of the lethal allele, q . This is an objective measure of the number of affected fetuses produced and hence a measure of the problem which needs to be addressed. How this burden is interpreted remains subjective and depends on the observer. If all affected embryos could be identified and terminated *in utero*, then some observers would then regard the burden as having declined from $q = 3\mu$ (equation 3, assuming equal mutation rates in both sexes) to zero. Alternatively, those ethically opposed to termination may regard its magnitude as

having increased from zero (genetic disease perhaps being a manifestation of divine will) to the number of fetuses at risk from termination, i.e. $q = 4\mu$ (equation 4, assuming equal mutation rates in both sexes). Alternatively, if medical intervention can ameliorate the disease, such as haemophilia, the increasing burden becomes economic rather than social. There is no objective resolution to this problem of measurement; however, it is interesting to note that of the three technologies considered here – embryo sexing, *in utero* diagnosis and family planning – it is the last which is most invidious. Application of the first two technologies raises the frequency of the mutation but prevent its expression (for example termination prevents birth of affected individuals), but family planning has no such remedial effect: it unambiguously increases the frequency of both the mutation and its associated disease. There is currently much interest in how changes in human society may have an impact on our genetic structure. Technological advances have reduced natural selection pressures in modern Western societies. The implication is that deleterious mutations will slowly increase and accumulate until a new equilibrium has been reached (Crow, 1997). This paper has described a much more immediate impact of reproductive technology: that the application of family planning and/or genetic screening may result in a large and rapid increase in the frequency of genes responsible for many of our more common and debilitating genetic diseases.

I thank anonymous referees for comments on the manuscript.

References

- Charlesworth, B. (1994). The evolution of lethals in the *t*-haplotype system of the mouse. *Proceedings of the Royal Society of London, Series B* **258**, 101–107.
- Conner, M. & Ferguson-Smith, M. (1997). *Essential Medical Genetics*. Oxford: Blackwell Science.
- Crow, J. F. (1986). *Basic Concepts in Population, Quantitative and Evolutionary Genetics*. New York: W. H. Freeman.
- Crow, J. F. (1997). The high spontaneous mutation rate: Is it a health risk? *Proceedings of the National Academy of Sciences of the USA* **94**, 8380–8386.
- Emery, A. E. H. (1991). Population frequencies on inherited neuromuscular diseases: a world survey. *Neuromuscular Disorders* **1**, 19–29.
- Emery, A. E. H. (1993). *Duchenne Muscular Dystrophy*. Oxford: Oxford University Press.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow: Longman.
- Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology* **12**, 12–45.
- Hartl, D. L. & Clark, A. G. (1997). *Principles of Population Genetics*, 2nd edn. Sunderland, MA: Sinauer Associates.
- Hastings, I. M. (2000). Models of human genetic disease: How biased are the standard formulae? *Genetical Research* **75**, 107–114.
- Hughes, M. I., Hicks, E. M., Nevin, N. C. & Patterson,

- V. H. (1996). The prevalence of inherited neuromuscular disease in Northern Ireland. *Neuromuscular Disorders* **6**, 69–73.
- Maynard-Smith, J. (1998). *Evolutionary Genetics*, 2nd edn. Oxford: Oxford University Press.
- Milunsky, A. (1998). *Genetic Disorders of the Fetus: Diagnosis, Prevention and Treatment*. Baltimore: Johns Hopkins University Press.
- Motulsky, A. G. (1997). Screening for genetic diseases. *New England Journal of Medicine* **336**, 1314–1316.
- Race, R. R. (1944). Some recent observations on the inheritance of blood groups. *British Medical Bulletin* **2**, 165.
- Short, R. V. (1994). Human reproduction in an evolutionary context. *Annals of the New York Academy of Sciences* **709**, 416–425.